Normative Defeaters

A constructivist approach[1]

Carla Bagnoli

1. The case for defeaters

Consider a rather ordinary piece of deliberation. I owe you a book, which I have borrowed.[2] I am under obligation to return it as I promised, but I have not read it, and I ask you if I can keep it until I do. You agree to let me have the book until I finish it. In this case, I am bound by the obligation to return the book as promised, but I am permitted to keep it for the time being. Suppose, again, that you decide to give the book to me. In this case, I am released from the obligation to return the book; or, better, there is no such obligation. In fact, returning the book would be discourteous or, at least, strange, because it would amount to rejecting a present. Then, suppose I discover that the book is not yours, and I am perplexed about what to do. We may disagree about whether I have an obligation to return the book to you as I promised, since you are not the legitimate owner. In this case, we disagree about the conditions under which this obligation binds.

How to construe these deliberative scenarios and understand their philosophical import is a matter of contest. The concept of "defeater" proves useful to articulate at least some of the subtleties of these cases. In the context of normative deliberation, a defeater is a normative claim that blocks, silences, undercuts, or undermines the normative force of other normative claims in rational deliberation. However, the defeasibility of normative claims names a problem, more than a

[2] This is an adaptation of the famous example about borrowed arms in Plato's *Republic*.

solution, and there are competing accounts of its philosophical importance. My proposal is that we take defeasibility to be an indication that a deliberative matter lacks normative determinacy.[3] This interpretation is broad enough to include cases in which normative claims fail to determine what to do (i) because there are stronger normative claims that override or cancel them,[4] or else (ii) because there are other kinds of normative claims that rule them out as irrelevant. In the former case, the normative failure is due to the presence of rebutting defeaters; in the latter, to undercutting defeaters.[5] Rebutting defeaters interact with other normative claims, by overriding, undermining,

---

[3] In some cases, the lack of normative determinacy is better understood as a failure of normative determinacy. In either case, the definition I offer is opposed to defeasibility as a failure of representation, see Redondo 2012. Redondo holds that moral realism, which considers moral reasons as facts, does not admit of defeaters: Either a consideration is a reason for something, or it is not. Redondo's characterization has the unfortunate consequence of ruling out moral realists who are in favor of particularism.

[4] I name only a few forms of normative failure, but the list is not meant to be exhaustive, see the pioneering work on the subject in Nozick 1968.

[5] For an informal definition of rebutting and undercutting defeaters, see Horty 2014, pp. 50-51. cf. Pollock 1999, p. 37. Horty distinguishes three accounts of reasoning by default rules. In the fixed priority theory, priority relations are fixed in advance, and all the defaults are universally applicable; in the variable priority default logic, the priority ordering is also defeasible; in the exclusionary default logic, default logic is used to determine whether certain default rules are to be excluded from consideration entirely. This latter case deploys a different sense of defeasibility, but I do not think the difference is as fundamental as suggested in Redondo 2012, p. 320. In any case, this difference does not affect my argument. A concern that I have about Horty's proposal is whether defeasibility may be preserved in a non-monotonic reasoning if the normative relations among reasons and defeaters are governed by fixed priority rules. I will consider the implications of

cancelling, or weakening them. Undercutting defeaters, instead, do not interfere with other normative claims: They exclude them from considerations entirely so as to deny their normative status.[6] I choose to formulate defeasibility in terms of generic normative claims, rather than in terms of reasons or obligations, because it is an open question whether defeasibility is a property of reasons or obligations.

Understood as a normative failure, defeasibility may be defended in two ways. *Moderate* defeasibility is the view that the same normative claims differ in normative status and moral valence according to deliberative contexts. However, this is not to say that the normative status and moral valence of such normative claims depends on the background contexts. *Radical* defeasibility is the view that moral reasons are considerations that count in favor of doing something in specific contexts, but not in others; hence they are particular.[7][8] Whether and how these two sorts of defeasibility are related is a controversial matter. They both seem to imply that knowing what to do is a sort of practical knowledge that cannot be codified. However, radical defeasibility denies that moral principles may be the source of practical knowledge and play an active role in practical

---

admitting moral defeaters for moral authority, without entering the normative debate about which kinds of defeaters, and how they are logically organized. These issues pertain to a formal account of the normative theory of reasons, which is beyond the scope of this essay.

[6] Among undercutting defeaters there are defeaters that show how some kinds of reasons do not apply in some contexts, because they are the wrong kind of reason.

[7] This is Dancy's position: "It is not as if it is some reason for me to believe that there is something red before me, though that reason is overwhelmed by contrary reasons. *It is no longer any reason at all* to believe that there is something red before me; indeed it is a reason for believing the opposite" Dancy 2004, p. 74. See also Dancy 1993, pp. 84-86, Wiggins 1993, Horty 2014, p. 152.

[8] Particularists often recognize that broad generalities are relevant e.g. in moral education and conversation, see e.g. Little 2000.

reasoning, while moderate defeasibility does not commit to this view. Radical defeasibility is defended on the basis of the thesis that reasons are holistic, but it is controversial whether this is a sufficient basis for denying that moral principles convey practical knowledge about what to do.[9] Finally, holism about reasons may not be the only source of defeasibility. In fact, the moderate view accommodates important cases of defeasibility, which are not due to holism: Cases in which the normative significance of some claims does not change because of a change in the background context, but because of the agent's change in view. This is what we ordinarily call making progress.

My aim here is not to offer a complete theory of defeasibility, but to show that the problem of defeasibility hinges on a more fundamental issue about the source of authority of moral reasons. I will argue the legitimate source of moral authority is rationality, and that rational justification deploys universal principles. However, normative defeasibility raises important issues about the possibility of practical inferences. I will argue for a version of Kantian constructivism that provides an alternative view of practical reasoning, where principles are both structural and generative. On the basis of this account of reasoning, I defend the claim that the deontic importance of moral considerations may survive the fact that they are defeated in deliberation.

2. Defeaters as a moral problem

We all routinely experience situations where a consideration that counted as a reason for doing something in one context fails to provide us with a decisive reason in another. Is this because reasons necessarily refer to particular agents in their particular circumstances? This question

---

[9] For an argument that justifies moral particularism on the basis of holism about general reasons, Dancy 1993, pp. 84-86, Wiggins 1993, Dancy 2006, p. 7, Lance & Little 2008. For Recent arguments that dispute the relation between holism and particularism see Millgram 2005, Väyrynen 2006, Väyrynen 2009, MacKeaver & Ridge 2005, MacKeaver & Ridge 2006, Schroeder 2011, Darwall 2013.

becomes more poignant when we reconsider the example of the borrowed book. This is not because it is an example about moral obligations, and moral obligations enjoy a special normative status or concern especially serious matters.[10] Rather, this is because the concept "obligation" is a deontic concept that differs significantly from the concept of "reason" understood as "considerations that count in favor of something."[11] This paper is an attempt to relocate the debate about defeaters in a larger debate about the legitimate source of moral authority, by bringing this distinction to the forefront.

Here is the predicament that needs to be addressed. As deontic operators, moral obligations are constraints: Their job is to limit what we do. Whether constraints are morally legitimate importantly depends on their rational justification. A strong argument about moral legitimacy establishes that moral claims should be rationally justified on the basis of principles that are universally binding for all relevant agents.[12] This requirement expresses the (moral) preoccupation to avoid arbitrariness, discrimination, and is meant to block exceptions in one's favor. Take the case of a norm regulating access to the library, which says that users who do not return books should pay a fine. Most of you would be surprised to learn that the norm is applied to everyone but Barbara; you would ask for an explanation. Most of you would object if the librarian accounts for the exception by explaining that Barbara is the daughter of a generous donor. This exception can and

---

[10] I share Dancy's view that moral reasons have the same status as other normative reasons, see Dancy 2006, p. 132. I do not argue that moral obligations differ from other sorts of obligations because of their special content or status.

[11] This expression is Scanlon 1998, p. 17. But it has become current in debates about reasons, and routinely deployed as a primitive (that is, not further analyzable) term in practical reasoning.

[12] I take "relevant agents" here to mean agents that are capable of understanding what moral obligations operate and how they applied to them. In short, this means agents capable of responsibility. For the purpose of this argument, further qualifications are not necessary.

should be objected to as unfair and discriminatory. The requirement of universality serves the purpose of excluding such exceptions as immoral because it is arbitrary.

On the other hand, a widely shared view is that the appropriate moral recognition of distinctive and specific moral aspects of contexts requires granting exceptions.[13] Suppose that a regular theater ticket is $30. Anyone who wants to go to the theater ought to pay the ticket. Some, if not most, of you would judge the usher who makes an exception for Andrej because he is poor as thoughtful and considerate. Perhaps, some would object to the institutionalization of the exception, e.g. a principle establishing that all certified poor people could enter theaters without tickets, on the ground that such a principle would be ultimately self-defeating, e.g. causing theaters to close down. More importantly, the usher's exception may resist the formulation of any more complex principle (even though inspired by equity). In fact, the case might be such that the usher intends to make an exception for Andrej, and only for Andrej, out of gratitude for Andrej's dedication as a former director of the theater on 42nd Street. How do we account for this exception, and avoid immoral arbitrariness?

Arguably, any theory of reasons for action should take into account special circumstances, but defenders of radical defeasibility demand that deliberators pay attention to particulars in ways that make universal principles inapplicable. They claim that the moral valence of reasons varies according to changes in the background circumstances. Thus, there is an interesting disagreement about how to explain cases such as Andrej's. According to some, the explanation of the exception must be principled, or else reasons become arbitrary. According to others, instead, to search for a principled account of exceptions is to misunderstand the very problem of defeasible reasons. Notably, for Jonathan Dancy, the holistic nature of reasons counts against the role of moral

---

[13] On the force of this argumentative move, see also Holton 2002, and Millgram 2005.

principles. This is controversial.[14] Dancy acknowledges that there is a logical gap between holism about normative reasons and moral particularism (Dancy 2006, pp. 81-82); and he also acknowledges that for some normative subjects (such as mathematics) universal principles play a significant role (Dancy 2006, p. 82), but he denies that the same holds true in morality, insisting that holism about normative reasons leads to moral particularism (Dancy 2006, p. 132).[15]

Like Dancy, I think that moral claims do not behave differently than their non-moral counterparts. I also agree that the recognition of variability in the normative significance and moral valence of claims across contexts urges us to reconsider the alleged form of practical reasoning. And, like the particularist, I think that the phenomena of variability described above forces us to reject a classical account of practical inference. However, in contrast to Dancy and other particularists, I will argue that a radical view of defeasibility mischaracterizes and misunderstands the basic problem of the variability of normative claims, and commits to a problematic account of moral authority. This is because they equate the moral authority and deontic importance of reasons with their overridingness in deliberation. In the next section, I begin by situating normative defeasibility in its proper context.

3. Normative defeasibility in context

The issue of defeasibility arises in the context of the rule of law. In a pioneering work on defeasibility, H.L.A. Hart argues that for normative concepts such as "contract" it is theoretically impossible to identify necessary and sufficient conditions of application because it is impossible to

---

[14] For instance, McKeever & Ridge 2005, McKeever & Ridge 2006 argue that holism presupposes generalism. On the claim that reason holism is compatible with generalism, see also Holton 2002, Schroeder 2011, Horty 2014, p. 153.

[15] I suppose this is what Millgram calls "the defusing move," not quite an argument, Millgram 2005.

specify *ex ante* (viz. before the application to particular cases) all the implicit exceptions. It is always possible to discover new exceptions to the rule that we have not anticipated, and this openness precludes the final delimitation of necessary and sufficient conditions. To ignore the irreducibly defeasible character of normative concepts is to misrepresent them (Hart 1955, p. 150). This semantic claim about the open-texture nature of normative concepts and the normative claim about the defeasibility of legal claims are distinct, but they can be regarded as broadly analogous, in that they both have the effects of creating the condition for discretional judgment.[16]

It is noteworthy that the reference to the discretionality of the judge is internal to the rule of law, e.g. as opposed to ruling by decree. That is, legal order is afforded by a principled, law-like structure, and the spirit of the law is an important substantive criterion for determining the admissibility of defeaters. The judge is required to consider whether the newly discovered exception can be absorbed as an exception to the rule without affecting the basic meaning of the rule. Furthermore, the principled structure of law is conceptually related to legitimacy. Indeed, it is a basic condition for its legitimacy. It is a conceptual matter that legal obligations imply universal principles because they apply in a shared context of mutual accountability and answerability.

Arguably, this feature of legal defeasibility is broadly analogous to moral defeasibility, at

---

[16] See Helm 1968. To some extent, the claim is analogous to the claim about the shapelessness of the moral in respect to the natural. The "shape" of a concept refers to the real similarity between all, and only the things that fall under it. The claim admits to different formulations, but it is generally invoked to explain the phenomenon that the extensions of evaluative terms and concepts seem to outrun non-evaluative classifications, see Kirchin 2010. For instance, McDowell holds that there is no guarantee that the extensions of evaluative terms or concepts are "shapely" in the sense that they correspond to kinds into which things can in principle be seen to fall independently of an evaluative outlook," McDowell 1981, p. 216. Dancy 1993, pp. 84-6; Wiggins 1993; Little 2000; Kirchin 2010.

least, as far as moral obligations are concerned.[17] The purpose of this section is to establish two points, which will be further developed in the course of the argument. First, that some important phenomena associated with defeasibility make sense within a principled structure. Second, that the principled structure of normative claims is a condition of their legitimacy. These considerations suffice to show that the radical view of defeasibility, which disposes of moral principles on the grounds of variability of normative claims across contexts, mischaracterizes the moral problem of defeaters as presented in section 1. As a consequence, it disregards the fundamental role of principles at the level of justification. This is a risky position to take, because it may show lack of appreciation of principled reactions against discrimination and arbitrariness.

These remarks may seem to be based on a broad similarity between ethics and law, regarding the form of justification. An objection may be made that this is exactly what is at stake. However, my case for the moderate view of defeasibility does not rest on an unqualified analogy with law. In fact, I recognize a seeming asymmetry between the moral and legal case. In the case of legal defeasibility, the problem is resolved by appealing to the judge as the institutional figure that exercises her authority *ex jure*.[18] In the moral case, the issue of discretional judgment is complicated by the fact that there is no institutional source of moral authority, or at least claims that morality is merely an institution, hence thoroughly conventional, is controversial.[19] Thus, the disagreement

---

[17] For instance, Stephen Darwall regards this analogy as limited to the concept of moral obligation, but recognizes an asymmetry between legal and moral contexts, see Darwall 2013, pp. 172-73.

[18] I leave aside the complexity of the legal case. There is a disagreement about the source of legal authority. In contrast to legal positivism, some argue that the legitimacy of law must refer to moral principles. Furthermore, in legal contexts there is a further distinction to draw between principles and rules, see Redondo 2005.

[19] It is undeniable that there are institutional and conventional aspects of morality; the issue is whether moral obligations rest solely on convention. See e.g. Williams 1985.

about the philosophical significance of moral defeaters revolves around a deeper disagreement about the status of moral claims and the source of their legitimate authority. My point will be that the source of authority of moral claims is the very same as the source of authority of normative claims, that is, reasoning.

4. Oughts and obligations

A standard strategy to account for the phenomena of variability of moral claims across contexts is to deploy the term *prima facie* ought, which stands for a defeasible duty. By introducing this terminology, W.D. Ross clarifies that the qualification *prima facie* does not mean that the duty is only apparent; instead, it is "an objective fact involved in the nature of the situation". [20] Indeed, to capture the philosophical significance of cases where moral duties conflict, it is crucial to take into account that the "ought" not acted upon does not lose its moral and deontic importance. Part of the ordinary understanding of moral duties is that they remain normative and significant even when they are occasionally defeated by other duties, as it happens in the case of moral conflict that – happily or tragically – allow for a resolution. In such cases, the issue arises about the relation between defeaters and defeated duties.[21]

---

[20] Ross 1930, p. 20. I believe that it is a mistake to conflate the nature and normative status of the all-things-considered judgment and the normative status of *prima facie* duties that are thereby defeated; see also Williams 1963, pp. 184-185.

[21] Williams criticizes ethical theories for failing to account for regret, "basically because they eliminate from the scene the 'ought not acted upon'," Williams 1963, p. 175. The thought is pursued in Dancy 1993, chapter 7. By contrast, for Hare *prima facie* duties are defeasible rules, which can be overridden in deliberation, without leaving a moral remainder. Since such rules are associated to deontic emotions (e.g. guilt and regret), this explains why such emotions are present even after the moral conflict is resolved, Hare 1981, p. 39. For an account of moral residue in a balancing

Current debates about defeasibility build upon some diagnosed failures of ethical theories to plausibly account for the aftermath of moral conflicts. The discussion is complicated by the fact that the term "ought" is often understood in a broad sense, which includes unqualified and qualified duties, obligations, and reasons (Dancy 1993, pp. 31-37; Horty 2014; Broome 2004). There is a rationale for this choice of term, even though it is not always made explicit. The broad sense of "ought" is instrumental in any theory of practical reasoning that aims to guide ordinary agents by providing them with the means for calculation. The broad sense of ought serves as a current term for aggregating reasons for action. This is apparent in the case of the "balancing structure" model of reasoning.[22] On this view, the agent is supposed to select the option whose overall weight is greatest. Take, for instance, John Broome:

> Each reason is associated with a metaphorical weight. This weight need not be anything
> so precise as a number; it may be an entity of some vaguer sort. The reasons for you to
> ϕ and those for you not to ϕ are aggregated or weighed together in some way. The
> aggregate is some function of the weights of the individual reasons. The function may
> not be simply additive. […] It may be a complicated function, and the specific nature of
> the reasons may influence it. Finally, the aggregate comes out in favor of your ϕing, and
> that is why you ought to ϕ. (Broome 2004, p. 37)[23]

---

structure, see Nozick 1981, pp. 489. See Bagnoli 2007.

[22] I consider a very simple example of balancing structure, but things get complicated very soon: See Nozick 1968, Nozick 1981, 479-494. Despite the elegance and depth of Nozick's account, I think that my objection applies, which is that moral obligation resists this treatment.

[23] Broome's balancing structure does not use numbers, and thus it is not additive. However, it is not clear to me by what more "sophisticated function" it generates aggregates. I suspect that there is no answer to this question, but reservations concern the overall project, which aims to determine action by producing aggregates of reasons. I believe it is misguided. On measurement of moral weight, see

It is not at all obvious that the concept of moral obligation is reducible to the concept of what there is most reason to do. Obligations refer to requirements that would be wrong for the agent to violate. Suppose you have an obligation of fidelity to tell the truth to your friend, but you also have an obligation to protect her from some horrifying truths. Upon deliberation, you decide that there are stronger reasons not to be straightforward with your friend. It is predictable and commendable, not only conceivable, that you may still feel bound by the obligation of fidelity, even though you have convinced that you have done the right thing. This is to say that the deontic importance of moral obligation survives the fact that it has been defeated in deliberation. I leave aside some complications about the nature of the remainders of correct moral deliberation. The example suffices to establish that there is something about the concept of moral obligation that is not preserved by the analysis in terms of reasons.

To vindicate this claim, there are at least two strategies. The first is to argue that the concept of moral obligation differs from other normative concepts, such as reasons. For instance, one may argue that in order to hold people answerable for complying with moral obligations, we have to assume that they can know that they are obligated, that they can regulate their conduct by this public knowledge. [24] A second strategy is to argue that rational justification necessarily appeals to principles, and ultimately amounts to a public justification, in ways that render the balancing model inadequate as a form of practical reasoning. I will adopt the second strategy, and in the remaining part of the essay I will defend a constructivist account of rational justification.

5. Against applied principles

Nozick 1981, pp. 490-492.

[24] This view is defended in Darwall 2013. In section 9, I will briefly comment on this strategy, and explain why the constructivist account I propose is preferable.

Kantian constructivism is the view that there are objective criteria for deciding what to do and that such criteria are principles of action that are both subjectively authoritative and universally binding.[25] The qualifier "Kantian" signals that the constructivist theory I defend takes reasoning to be autonomous, that is, independent of what reasoning is about, and of what reasoners happen to desire or value. The function of practical reasoning is not to track moral truths that are there prior to and independently of its activity. Thus, the criteria of correctness for practical reasoning do not conform to its object domain. An important implication of this view is that there is no definite subject domain, that is, there are no objects of reasoning prior to and independently of the very activity of reasoning.[26]

A prominent argument in support of constructivism establishes a distinction between reasoning aimed to action, and reasoning aimed at knowledge. According to Christine Korsgaard, acting on and grasping a bit of knowledge differ, and the difference is such that it counts against the conceiving of moral principles as conveying knowledge about what to do.[27] She names this the

---

[25] Kantian constructivism names a rather diversified cluster of theories, which differ as to the scope and aims of construction. See Bagnoli 2011a.

[26] See O'Neill 1996, p. 85. This claim has been often defended in proceduralist terms, see e.g. "The procedural moral realist thinks that there are answers to moral questions *because* there are correct procedures for arriving at them. But the substantive moral realist thinks that there are correct procedures for answering moral questions *because* there are moral truths or facts," Korsgaard 1996, pp. 36-37.

[27] Korsgaard's argument against the application model builds upon two distinct claims. First, the claim that cognitions of a piece of reality cannot be "compelling:" "For think how that account would have to work. The agent would have to recognize it, as some sort of eternal normative verity, that it is good to take the means to his ends. How is this verity supposed to motivate him?" Korsgaard 2003, p. 110; Korsgaard 1996, pp. 16, 38-40. Second, the claim that for whatever moral

application model of practical knowledge. When moral principles are taken to state some piece of knowledge, the question arises how to put that piece of knowledge into practice. According to Korsgaard, the question can be settled only by introducing a new set of principles, about which the same issue arises. The application model is thus shown to face a regress.[28] But the fundamental mistake of the application model of moral knowledge is that it misrepresents the relation between judgment and action, by treating moral principles as stating moral facts. The ability to apply knowledge in action presupposes the capacity for action, but this is exactly what we are trying to understand. She concludes that: "The model of applied knowledge does not correctly capture the relation between the normative standards to which action is subject and the deliberative process" (Korsgaard 2003, p. 112).

Thus far, the constructivist argument is not hostile to defeasibility. In fact, it carries some force against traditional models of practical reasoning, which identify the role of moral principles as the premises of a practical inference. In contrast to such models, constructivism denies that moral

---

principle stating what to do, the question arises whether and how it applies. "If it is just a fact that a certain action would be good, a fact that you might or might not apply in deliberation, then it seems to be an open question whether you *should* apply it," Korsgaard 2003, p. 112, Korsgaard 1996, pp. 44-47. For a critique of both claims see Bagnoli 2013. For present purposes, only the second claim is relevant.

[28] Korsgaard's argument draws on an analogy with maps: "If to have knowledge is to have a map of the world, then to be able to act well is to be able to decide where to go and to follow the map in going there. The ability to act is something like the ability to *use* the map, and that ability cannot be given by *another map*. […] goodness in action cannot just be a matter of applying our knowledge of the good—not even a matter of applying our knowledge of what makes action itself good. This is because the ability to apply knowledge *presupposes* the ability to act," Korsgaard 2003, p. 110. The analogy is problematic on its own, but this is a side issue here.

reasoning deploys subsumption or deduction. The role of moral principles must be conceived differently. Korsgaard presents the difference by taking practical reasoning to be a decision procedure. I think this is a misleading formulation of constructivism, which does not fully capture the real difference between classical and constructivist accounts of rational justification.[29] At this stage of the argument, however, there is another live option beside constructivism. Korsgaard's argument about the distinction between grasping a bit of knowledge and acting does not establish that there is no room for practical knowledge. In fact, what the argument seems to show is that moral principles do not play a pivotal role in determining what to do. The matter may be decided on judgment.

This is why Mark Lebar writes: "The motivation for being constructivists should move us to be particularists as well" (Lebar 2013, p. 200). In his view the unintended result of Korsgaard's argument is to expose the inefficacy of moral principles (Lebar 2013, p. 192). The inefficacy of moral principles shows that moral principles fail to determine what to do, because they do not contain specifications about the conditions of their applicability.[30] The argument in support of this radical view of defeasibility is that any given consideration can enter into deliberation with different kinds of normative force in different conditions. The authority, the practical importance, and the moral valence are all normative aspects of reasons that vary across different contexts, and inevitably elude moral principles. As anticipated in section 2, there is a missing step in this argument for radical defeasibility; in short, holism does not definitely demonstrate moral particularism. But I am interested in Lebar's diagnosis for the Kantian failure to treat defeasibility. In his view, Korsgaard's

---

[29] On a more robust conception of construction, constructivism qualifies as a distinct form of rational justification, which is alternative to practical inference.

[30] "Deciding which cases are which is, of course, a deliberative problem and that is an issue on which principles give no guidance at all. Principles tell us nothing about where or when they are to be applied," Lebar 2013, p. 192.

argument against the application model shows that we must do without principles because they are too abstract to be useful in identifying moral reasons and their defeaters.

5. Provisional principles as default rules

Thus put, the problem of moral defeasibility is pressing but not new. The challenge for Kantian constructivists is to show how universal principles can effectively guide human agents, given the heterogeneous mass of considerations that are morally relevant, and which they cannot anticipate. A standard strategy to address the problem is to invoke the distinction between ideal and non-ideal conditions of rational agency. While practical laws are binding for all rational beings insofar as they are rational, they are not immediately accessible to human agents because of the embodied aspects of their agency. Human agency is subjected to practical and epistemic limitations. Because of these limitations humans cannot grasp what they ought to do. Reasoning is the activity in which they engage in order to figure out what to do, and it is an activity governed by principles.

However, there is a question about the level of universality that these principles require. Korsgaard introduces a distinction between "merely universal principles" and "provisional principles:"

> There is not reason to suppose that we can think everything in advance. When we adopt a maxim as universal law, we know that there might be cases, cases we had not thought of, which would show us that it is not universal after all. In that case, we can allow for exceptions. But so long as the commitment to revise in the face of exceptions is in place, the maxim is not merely general. It is provisionally universal. (Korsgaard 2009, pp. 74-75)

Provisional universality importantly differs from generality, which offers a trivial explanation of the apparent defeasibility in ordinary deliberation because generalizations obviously admit of exceptions. Korsgaard explains the relevant difference between generality and provisional universality in terms of the agent's commitment to revise the provisional principle whenever she

encounters an exception. If the principle of action is thought to be general, the presence of exceptions is to be expected and does not force the agent to revise the principle. By contrast, if the agent is required to act on a maxim that she can regard as a practical law, and she faces what appears to be an exception to that practical law, then she is bound to revise her principle of action. That is to say, the agent who is committed to act on the basis of reasons that count as practical laws is thereby committed to either abandon or revise her subjective principles when they do not meet such a requirement. The revision is meant to make it more specific and determinate, by including the exception in its formulation (Korsgaard 2009, pp. 73-74).

This argument is motivated by the recognition of a weakness in the Kantian appeal to universal principles. Thus, Korsgaard appears to agree with her critics that ordinary moral thinking appeals to moral principles understood as "summary rules" because strictly universal principles are unavailable to human agents.[31] That is, Korsgaard concedes to critics that ordinary moral thinking is governed by norms that cannot attain strict universality. I think this is a mistake, which bears very high costs.

On the model that exploits two senses of universality, the genuine deliberative work appears to be done at the level of revisions. At this level, the agent is supposed to consider that the moral principle does not apply, then is supposed to identify the moral defeater, and then decide whether and how the defeater is to be incorporated into the principle. Supposedly, some exceptions can be built into the principle so as to make it more determined, focused, specific, and thus narrower in content. However, it might still be theoretically impossible to anticipate all possible defeaters and incorporate them in the formulation of moral principles. In any case, this is an extraordinary amount

---

[31] This is Richard Hare's argument in support of *prima facie* duties at the ordinary level of moral thinking, see Hare 1981, chapter 1. The term "summary rules" is introduced by John Rawls in contrast to the "practice" conception of principles, see Rawls 1955.

of work for an agent to accomplish in any deliberative context.[32] By modifying the content of the principles so as to accommodate relevant exceptions, we end up with entities too complex to be manageable and practical. Furthermore, it is likely that such sophisticated principles would appear to have little moral authority, since agents would find them hard to learn and to be guided by. In sum, Korsgaard's distinction introduces a complication that does not solve the problem of defeasibility and brings us back where we started.

A second set of considerations reveals that Korsgaard's amendment is not congruent with a Kantian theory of rational justification. By instituting the distinction in universality, Korsgaard treats provisionally universal principles as summary rules that the agent takes as valid by default, until she encounters an exception. Here is an example: Hannah thinks that all passengers ought to pay the train ticket, but today there is a protest against the railway company, and she decides that this is an exception to take into account. Then, she revises the principle accordingly: "Passengers ought to pay the train ticket, except when they take part in a protest against the railway company." This principle guides the agent in particular occasions in which she deliberates about paying the train ticket. In all similar cases in which Hannah rides on the train, and happens to consider whether or not to pay the ticket, she relies on this principle. But she does not rely on the principle because she knows that it is valid. In other words, she treats the principle as a default rule, which can be defeated by stronger reasons. The question to ask is on which authority the agent operates.

In support of relying on default norms, it is often said that these norms spare the agent the worry to deliberate all over again in any new deliberative scenario. However, this is a risky position for Kantian theorists to take. First of all, this approach dispenses the agent from looking harder and

---

[32] Millgram also argues that the strategy of institutionalizing exceptions produces perverse results, insofar as it raises the complexity of the principles of action. As already noted by Hare and Nozick, complex principles are not only hard to formulate, but also difficult to learn. Their efficacy is doubtful.

may encourage a passive reliance, insofar as it legitimizes principles established by habit or tradition.[33] The appeal to moral principles that are established prior to and independently of deliberation goes against the requirement of autonomy, which is a defining feature of the Kantian theory of practical reasoning.

Indeed, the emerging picture of deliberation as governed by default norms appears to be at odds with the general Kantian view of reasoning as a constructive activity, governed by the requirement of strict universality.[34] Upon detection of the exception, Korsgaard advises, "We must now go back and revise it, bringing it a little closer to the absolute universality to which provisional universality essentially aspires" (Korsgaard 2009, p. 74). Thus implemented, the universality test does not check whether the particular subjective maxim that the agent initially endorses coheres with its strict universalization. Rather, Korsgaard's account of revision resembles a step-by-step attempt to approximate a regulative ideal. The difference in the practical import of these two tests of coherence may be illustrated as follows.

Consider the case in which a moral agent relies on a default moral principle, which proves to be reliable, in standard circumstances. For instance, Martha knows that it is a good norm to return borrowed books to their legitimate owners. Using the distinction between strict and provisional universality, allows us to say that Martha has not encountered any exception. The absence of registered exceptions does not justify Martha in thinking that the moral principle holds without

---

[33] This is one of Dancy's objections against reasoning by principles, see Dancy 2004, p. 64. For a critique of this position see Bagnoli 2011b

[34] I take much for granted here. For a thorough discussion of universality as the criterion of rational knowledge and action, see Engstrom 2009. I agree with much of the argument that Engstrom offers in support of situating Kant's appeal to the universal law in the tradition of practical cognitivism. I also share Engstrom's view that a constructivist account of practical reason serves the purpose of vindicating the objectivity of practical knowledge, see Engstrom 2013.

exception. It might be that she happens to rely on a principle that has been successfully applied in all contexts she has tried thus far. Or, worse, it could be that Martha is not sufficiently sensitive, attentive, or competent so as to recognize an exception when she sees one. It might also be that she is so often distracted that she rarely identifies the proper scenario in which the moral principle standardly applies. As things stand, Martha might just be fortunate. How would she know? In a system regulated by default norms this question is irrelevant, if not unintelligible.

Given human constitutive, epistemic, and practical limitations, it is hard to imagine how human agents are supposed to attend at the revision of their subjective principles of action, build the exception into the provisionally universal principle, and proceed to make them strictly universal. People are not able to construct such principles on purpose; if they happen to formulate such principles they would not even know why. This is because the gap between provisional universality and strict universality may be such that humans cannot calculate it or determine it. Whether provisional principles happen to meet the criterion of strict universality and qualify as natural laws, it is by chance as far as human agents know. By contrast, the very purpose of the test of strict universality is to rule out maxims (reasons for action) that are merely fortunate. By proposing that agents modify their initial principles so as to make them increasingly complex, one is making them so rich in content that they cannot be fully known and managed. This fact does not merely show the severe limitations of our epistemic powers; rather, it also calls into question the source of authority of moral reasons. The fact that one cannot know on which principle one is acting upon is a serious moral problem. In fact, it is precisely the problem that Kantian theory is designed to identify by drawing the line between merely fortunate and morally valuable actions. Korsgaard's proposal threatens to lose what is distinctive to the Kantian approach to principles, that is, the claim that they are the very source of genuine authority.

The difference between provisionally and strictly universal principles is not marginal as Korsgaard believes. I hope to have shown that her argument confuses two distinct ways of reasoning about what to do: The first appeals to default norms, the second one appeals to universal

principles. To disregard this difference makes it difficult to understand what the normative status of the subjective principle of action is: Why does the agent formulate the subjective principle to act, how does she recognize a defeater, and on which authority does she go back and revise the principle? There seem to be at least three competing answers on the table: by judgment, by default, and by principle.

6. Defeaters and exceptions

On the Kantian view, the only source of normative authority is reasoning. The activity of reasoning is principled, hence governed by the requirement of universality. It follows that, strictly speaking, there can be *no exceptions* in a social world governed by Kantian practical reasoning. Of course, it happens all too often that particular agents reason insufficiently, commit mistakes, and thus act on the basis of maxims that are not principled. Such agents may find themselves in circumstances where their subjective principles of action lead to unwelcome results (e.g. they cannot borrow books from the public library because they have not returned many of them), and thus they are forced by the circumstances to reconsider the ground of action: This operation amounts to re-describing the candidate action. To preserve the practical significance of moral principles and their authority, the agent must think again on which grounds they act. This procedure is radically and importantly different from admitting an exception to the principle. The fact that the agent experiences the need to rethink the grounds of her action is a morally legitimate concern only insofar as her reasoning is principled. Otherwise, the agent would be at liberty to give up a moral principle, whenever she wants. It seems that the only way one can use reasoning to rethink our reasons is by considering exceptions to the principles; and this operation cannot be conceived as an external attack to the principles' moral authority.

The operation of "making an exception" does not require the agent to rethink the principle all over.[35] In fact, it appears that the moral and practical significance of exceptions is such that it presupposes the principle of which it is an exception. Free riders make exceptions in their favor, and their action is successful exactly because all/most others act in compliance with the principle. To admit that Andrej enters the theater without paying for the ticket counts as an exception on the understanding that people normally pay. That is, the practice of granting exceptions relies on the normativity of universal principles of which they count as exceptions.[36]

How are exceptions important? Elijah Millgram argues that making exceptions is a necessary precondition of successful agency. [37] This is a bold claim, which directly calls into question any principled view of rational agency. Millgram is right that making self-serving exceptions may be the only condition for an agent in particular circumstances to exercise his agency successfully (i.e. to act rather than to attempt at doing something). However, I do not see that his argument establishes what he claims, i.e. that making self-serving exceptions is a necessary

---

[35] Elijah Millgram forcefully makes the point: "Making an exception and building an exception-shaped twiddle into a rule work differently in a social world," Millgram 2005, p. 103.

[36] One may object that there are established practices of "institutionalizing exceptions," e.g. "Theatre tickets are $5, except for senior citizens." However, in this case, it is preferable to say that for the category citizen the rule does not range over the class of seniors citizens. The scope of application of principles may be broad or narrow (including large numbers of individuals or a singleton), but this is a different matter than the relation that principles bear to their exceptions.

[37] Millgram considers a fundamental flaw of Kantian theory that it cannot accommodate the need for exceptions, Millgram 2005, p. 108. He regards exceptions as "necessary preconditions for successful agency," Millgram 2005, chapter 5 §4.5. This is where Millgram and I part ways. He offers a theory of practical reasoning that is a viable alternative to Kantian constructivism, but I cannot take it into account here.

precondition of successful agency. Partly, the issue revolves around what we take successful agency to be. There are ordinary cases of deliberation in which rational agents prefer their immediate end of action to be frustrated, rather than violating the norm. In some of these ordinary cases, rational agents reason on the basis of considerations about the implications of his violation, because of their feeling of being obliged, or out of attachment for the symbolic value of the norm.[38] These are distinct varieties of normative behavior that do not require making exceptions for themselves, and yet they represent some sort of agential success, even though they involve the frustration of the intended end. Agential success does not amount to success in bringing about particular ends, because the agent may have reasons to have her local ends frustrated and, also, to describe the case in terms of the expression or realization of more significant values, rather than in terms of frustration. Indeed, this is an implication of treating some values as more important than others. The Kantian view explains the "need" to make self-serving exceptions as cases in which agents capable of rationality have not reasoned correctly, and thus they have not properly formulated their subjective principles of action, or else as cases in which their agency is obstructed.[39] The latter case is a matter of luck, and finds no remedy in the Kantian theory of reasons for action. The former, instead, is a serious matter, which we have to attend to next.

7. Rethinking action and making progress

In a Kantian theory of practical reasoning, the fact that a rational agent feels the need to make an exception to her principle requires her to reconsider her proposal for action. In this picture, the key deliberative work consists of rethinking the ground of action, rather than granting exceptions or revising default norms. Deliberation starts with a normative account of the circumstances of action,

---

[38] On the distinction between being "obligated," and being "obliged ," see Hart 1961, pp. 6-8.

[39] Obstruction may depend on the interference of others. This is a complex case, which would require special consideration, but I have to set it aside.

which may include large varieties of details, e.g. the perception of obstacles, the desire to realize some ends, or relate to others as independent sources of legitimate claims.

An important part of the deliberative problem is to identify the relevant description of what the agent has to do in the circumstances. Describing the circumstances and defining the action is the crucial step in practical reasoning that Kantian theory identifies with the formulation of a subjective maxim. It appears that the moral valence of actions will decisively depend on their descriptions. The crux is that actions indefinitely have many descriptions.[40] Thus, the moral valence of actions will decisively depend on the relevant description. This is known as the problem of relevant descriptions, and it is often presented as an argument against Kantian theories of practical reasoning.[41] Indeed, this is a problem that undermines deductive accounts of practical reasoning,

---

[40] The issue is "the amount of detail about the agents' circumstances and his proposed action which can be included in the maxim," O'Neill 1975, p. 37. A maxim is a practical principle underlying action in contrast to aspects of the action that are below the level of intention (sub-personal processes), and also in contrast to more specific intentions, see O'Neill 1989, pp. 129, 151-152, 158.

[41] To address the problem of the relevant descriptions, Barbara Herman argues that the CI-procedure applies psychologically to actual intention underlying the action, and produces a recognizable pattern of moral actions only insofar as the agents' psychology is relevantly similar. She stresses the psychological underpinning of the theory by introducing criteria of moral salience that help moral reasoning to be effective, see Herman 1993, chapt. 4, chapt. 3 §3, and chapt. 7 §4. There are two serious problems with Herman's proposal. First, on her account, the categorical imperative procedures constrain only agents who are already endowed with a moral sensibility and leaves the non-moral ones unconstrained. Second, as Millgram puts it, the "necessary but totally unaccounted for regimentation of agents' motivational structure has become the engine of the theory," Millgram 2005, p. 128 n. 39. For these reasons, Herman's solution to the problem does not

where the conclusion derives from fixed premises. This is the reason why philosophers impressed with the variance of moral reasons across contexts deny that moral principles play any significant epistemological role. Recall, for instance, W.D. Ross: "When I reflect on my own attitude toward particular acts, I seem to find that it is not by deduction but by direct insight that I see them to be right, or wrong. I never seem to be in the position of not seeing directly the rightness of a particular act of kindness, for instance, and of having to read this off a general principle – 'all acts of kindness are right, and therefore this must be, though I cannot see its rightness directly'."[42]

One might object that the dismissal of deductivism is too quick, insofar as ordinary inferences can be supplemented with default rules. Defenders of defeasible logic pursue this line. Others argue for the possibility and moral relevance of practical inferences, where the premises must contain implicit *ceteris paribus* clauses (Holton 2007). These philosophers maintain that every true moral conclusion follows deductively from a finite set of premises; and even when such premises are at least implicitly conditional, they are not trivial. These theories deal with the phenomena associated with defeasibility by conceding that moral obligations are always conditional.

By contrast, Kantian constructivism purports to address the problem of the variance of normative claims without renouncing the thesis that moral obligations are unconditional

---

seem to be congruent with the Kantian theory of practical reasoning, insofar as it is grounded in empirical psychology. The criteria of salience are a dubious basis for moral consensus and may undermine the autonomy of reason. For this very reason, Herman's account of practical reasoning does not qualify as constructivist, see Herman 1993, pp. 152-153. There is something to be said about the role of moral sensibility in practical reasoning, but it cannot be put in terms of criteria of moral salience.

[42] Ross 1939, p. 167. On this point, see Dancy 1993, p. 95. Lebar (2013) pursues this line, refreshing the often forgotten connection between Ross' intuitionism and Aristotelian ethics.

requirements of reason. Its basic claim is that the appeal to principles is key to explain how moral reasons remain authoritative when defeated, and also to distinguish between arbitrary and legitimate exceptions. Furthermore, the constructivist theory agrees with the particularist that the phenomenon of variability of reasons bears disastrous consequences for deductivist accounts of practical reasoning. But it argues that this diagnosis has a different impact. Defeasibility does not show that we can reason without principles, but that principles are not the fixed starting points of practical reasoning. The facts about defeasibility drive us toward rethinking the role of principles in practical reasoning.

8. Universality as the matrix of judgment

Reasoning is an activity that people do, and it is deeply connected to the sort of agents these people are. The structure of rational justification reflects the profile of the subjects that engage in such an activity. This is where the role of strict universality should be sought. The basic idea is that universalization is the form that best coheres with the profile of the practical subjects who engage in practical reasoning.[43] The argument in support of universalization is that this is a constitutive norm guiding the correct formation of reasons because (i) it takes practical subjects as independent and legitimate sources of valid claims, and (ii) it thereby allows finite and interdependent practical subjects to exchange reasons that they may find mutually authoritative in justifying their claims to one another. In other words, the basic argument for universality is practical: It warrants mutual intelligibility and accountability, which are crucial pre-conditions of cooperative and shared agency. Reasons function properly insofar as they are universally authoritative.

As a constitutive norm of rational agency, universality cannot possibly admit of defeaters. In the practice of rational justification, the job of the strictly universal principles is not to provide all

---

[43] Kant's argument for the categorical imperative is rather brief, see Kant Ak 5.19-30. Cf. Rawls 1980, Engstrom 2009, Engstrom 2013.

the relevant information that one can apply in practice so as to defuse all the possible defeaters, but to test the subjective maxims so that it can conform to a practical law.[44] To this extent, the requirement of universality is the matrix of judgment: It guides the rational agent in designing future action, by playing a structural role. It is *within* this context that the issue of defeasibility arises. In fact, in a Kantian constructivist account, any substantive moral reason is in principle defeasible. Substantive moral reasons are constructions that it is the agent's deliberative job to produce. In the paradigmatic case of rational deliberation, the agent is called to decide what to do, within the framework set by constitutive principles, but there is no guarantee that the agent always reaches a definite verdict about what to do.

Furthermore, it is conceivable that differently situated agents produce different reasons, despite the fact that they operate in relevantly similar contexts of choice. This common fact does not prove that the theory is false, nor does it show that the reasoners have failed the standard of reasoning, which would be paradoxical. As a matter of fact, there is no paradox to solve here. The problem lies in the ambiguity of the remarks about the "similarity of the contexts of choice." This phrase is taken to justify the presumption that the deliberator enters a fixed deliberative scene and is required to take account of the factual aspects of the context. But this is a misleading representation of rational deliberation. The mistake is to treat the factual aspects of the context as if they were morally relevant elements prior to and independently of practical reasoning. By contrast, the constructivist view is that the circumstances of action are not factual elements that are fixed prior and independently of practical reasoning. Instead, the delicate job of rational deliberation starts with describing what is the problem of acting in such circumstances. The moral problem is how to act in the circumstances.

---

[44] The requirement of autonomy excludes all thinking and acting that depend on "the contingent, subjective conditions that distinguish one rational being from another," Kant Ak 5.21.

This is the stage of construction where the phenomena of defeasibility take place. If I am right, the crucial step in reasoning about what to do is to formulate the subjective principles describing the intended actions. The requirement of universality serves as a test to discriminate actions that, at least some under some descriptions, can be rationally intended, and actions that do not. This is perfectly compatible with recognizing that the circumstances of actions were described differently, the agent would not have the reasons or the obligations she has.

A more general claim is that there are no objects of reason that qualify as such before the relevant subjects start reasoning. To say that *x* is an object of reason is itself a claim of reason. This means that only reasoning is entitled and equipped to identify its objects; to this extent, it is not determined by any pre-constituted objects, and it does not have a distinctive subject matter. There is no given, pre-fixed subject matter of reasoning that we access or represent by reasoning. Rather, the subject matter is constituted by the activity of reasoning. Indeed, this is one important sense in which reason is autonomous, hence productive and generative. The domain of reasoning is not discovered but built up by the activity of reasoning. The relation of reasoning to its objects is one of construction, rather than one of recognition or representation of a special realm of facts.

The autonomy of reasoning – that is, its independence from any prefabricated subject matter – reflects the autonomy of the relevant practical subjects. When reasoning merely conforms to external objects, its authority is conditional upon the agreement on such objects. This is why practical inferences can justify only conditional conclusions. By contrast, constructivist reasoning produces reasons that are genuinely authoritative, and guiding all relevant subjects, insofar as it is governed by structural universal principles. This is because constructivist reasoning does not start from points fixed by special communities, and thus it can legitimately claim universal authority.

The defeasibility of reasons is connected to another important aspect of reasoning. Undertaking reasoning is not only productive of normative objects, but also transformative for the subjects that practice it. This is to say that on a constructivist account, reasoning is practical not only because it effectively guides the agent in addressing the question of what to do, but also

because it affects the agent's general deliberative set, including her desires, interests, and concerns. This aspect of constructive reasoning vindicates the ordinary experience of rational deliberation, which is often perceived as an attempt to make progress in understanding what the situation requires. This represents an additional source of defeasibility that is not due to normative holism, but to the fact that rational agents are situated. They develop and deteriorate in time. Normative claims vary across contexts not because such contexts are arranged holistically, but because the agents describe them differently.

9. Reasoning as a transformative activity

To vindicate this claim, it is necessary to discard the view that practical reasoning is procedural and algorithmic, a merely formal scaffold by which we organize value-neutral factual matters. The sort of cognitive operations implicated in the application of a procedure or a function are mechanic and routinary, hence they do not presume any significant contribution from the agent. Of course, the agent may intervene at the very beginning, by picking the decision procedure, but for this choice to be reasoned rather than arbitrary we still need an account of reasoning. Reaching a formalization of a decision procedure is the result of reasoning, not the way to do it. I do not deny that it is possible and often rationally justified to construct and depend on procedures. What I dispute is the claim that practical reasoning amounts to applying a procedure.

On the basis of what I have argued thus far, I consider reasoning as an activity that produces a reasoned alteration, which may take different forms, e.g. a revision of belief, a change in concern, and also a change in view. Reasoning is something we do, an activity in which we engage because of our concerns, and starting from our perspective. The examples I presented show that emotional engagement may play a crucial role in generating a reasoned

change in view, and this is because emotions are distinctively perspectival.[45] Changes in view are perspectival. Nonetheless, the descriptions that one achieves as a result of focused emotional engagement can be more or less perspicuous and objective. In these final two sections, I provide a constructivist account of the "rational activity" of the mind, and show that such activity is not unconstrained, even though perspectival.[46]

To some extent, the constructivist conception of reasoning matches the definition of reasoning as "reasoned change in view" proposed by Gilbert Harman (1986).[47] Like Harman's, it starts with the acknowledgement that reasoning is something people *do.* Is this theory descriptive or normative, given that it deploys structural norms? Structural norms describe how we ought to think about what to do. The reasoning is productive or generative (of new intention to act/beliefs/judgments about what to do). Harman says it is both because it starts with what we do, but it also contains abstractions and idealizations that are necessarily normative.

Constructivism conceives of reasoning as an activity. Understood as the way in which reasoners produce reasoned changes in view, reasoning is not itself linear, and it is distinct

---

[45] This is feature of emotions that I consider decisive for the functioning of practical reasoning. See also Greenspan 1988. On the relation between objectivity and perspectival judgment, see Elgin 2010.

[46] The conclusion I reach in these sections are close the sort of constructivism that Caterine Z. Elgin defends about science: "If scientific facts can be relative and value-laden, there is a strong *prima facie* case for saying that relativity and value-ladeness do not undermine objectivity", Elgin 1997, p. 185. My argument will be on entirely different grounds, however, and it is to be situated in a larger picture of practical knowledge centered on autonomy, see Bagnoli 2013.

[47] The first programmatic chapter is interestingly titled "A Plea for the Study of Reasoning", Harman 1986, pp. 1 ff.

from some of its products, e.g. arguments, demonstrations, and proofs. The change in view is a complex perspectival change, which is the result of discursive and emotional intelligence and capacities.[48]

This conception of reasoning as an activity is in tension with the standard view that practical reasoning is either deductive or inductive, governed by the rules of logic, and takes the form of practical inference and syllogism.[49] Kantian ethics is often associated with deductive reasoning. As I have argued elsewhere, this interpretation is based on a misconception of the role that universal principles play in the Kantian account of practical knowledge. The constructivist account of practical reason has the advantage of clearing the ground from this mistake, by pointing to universality as the structure of rational activity, a mere matrix for practical judgment. Understood as the activity of revision, reasoning is constructive, that is, productive and generative. To ensure this constructive or practical function, reasoning must deploy principles merely as guidelines for constructing reasons,

---

[48] Harman concurs about the significance of emotions in changing view. See also Greenspan 1988, 2011.

[49] I have argued for this claim in Bagnoli 2016. Harman takes this stand but on the basis of different considerations: "Obviously, there is deductive argument, but it is not similarly obvious that there is deductive reasoning. Again, it is not clear that there is such a thing as inductive reasoning […] Analogous remarks also apply to the suggestion that there is such a thing as practical syllogism", Anscombe 1957, p. 57. A syllogism is a form of argument, and although there is practical reasoning, there is not obviously any such a thing as practical argument or logic and so not obviously any such thing as a practical syllogism […] It may be a mistake to expect principles of reasoning to take the form of a logic", Harman 1986, p. 6. Harman concludes that it is unclear how the formulation of logics of thought might contribute to the study of reasoned revision.

rather than as first points from which conclusions logically descend. [50] Rather than a formal procedure, a robust conception of construction is needed for constructivism to qualify as a distinct form of rational justification, genuinely alternative to practical inference and other standard logical forms of reasoning. On this robust conception, a construction of reason is productive or generative, that is, it produces something that was not in place before: a change in view. Such as change counts as objective insofar as it is produced by a rational activity constrained by universal principles.

This robust understanding of construction coheres with a more discriminating characterization of practical reasoning, which I take to be peculiar and distinctive of (Kantian) constructivism.[51] Activities differ from processes in that they require the active engagement of agents. Thus practical reasoning is an activity guided by structural principles of reason. In order to issue authoritative judgments, agents engaging in practical reasoning ought to be guided by such principles. A Kantian claim is that being guided by the principles of reason is a structural warrant of autonomy. This is to say that the sort of autonomy that reason provides is merely structural: it is a minimal requirement for being authoritative judges and actors. Of

---

[50] Constructivists themeselves have not been clear enough on the latter point. As a consequence, current formulations of constructivism do not fully capture the real difference between classical and constructivist accounts of rational justification. This is true of definitions of constructivism in terms of procedure, such as Korsgaard's, as well as of those characterizations that use the notion of entailment confined to practical standpoint, such as Street. It is not obvious that rules of inference, implication, and entailment have anything to do with practical reasoning. By contrast, see Bagnoli 2014, 2016.

[51] There are different varieties of constructivism and Kantian constructivism. I defend a cognitivist but irrealist account of constructivism, based on the claim of autonomy; for the latest statement, see Bagnoli 2013.

course, this minimal requirement does not guarantee that such reasoners are autonomous in more mundane and ordinary senses. More needs to be said about whether a given subject entering reasoning is autonomous. The definition above only fixes the condition for his activity to be autonomous, not a threshold for subjects to have personal autonomy.

The distinction between "activity" and "process" brings to light another difference in the way Harman and I conceive of reasoning. For Harman reasoning can be unconscious, while Kantian constructivism takes reasoning to be a conscious, self-reflective activity. This is not to deny that there are operations that we make unconsciously, and which may underlie our reasoning. However, that agents are conscious of their reasoning is a condition for their reasoning to be subjectively authoritative, that is, authoritative in the first person.[52]

Finally, the constructivist characterization allows for a finer distinction between practical and theoretical reasoning. Harman assumes that one's reasoning can be "separated into distinct segments of practical and theoretical reasoning", even though he recognizes that any given segment of reasoning can affect beliefs and intentions (Harman 1986, p. 1).[53] He also assumes that theoretical and practical reasoning concern beliefs and intentions, respectively; even though he recognizes that intentions may affect beliefs, and beliefs may affect intentions. By contrast, Kantian constructivism is interested in the distinction between

---

[52] Arguably, it is also a condition for moral responsibility and imputability, even though it might be that, in a different broad sense that does not include imputability, agents are morally responsible also for aspects of reasoning of which they are not aware, insofar as such aspects are expressive of character

[53] Harman also distinguishes between two phases of reasoning, the first consisting in the process of reflection, and the second consisting in the actual revision of one's view. Correspondingly, he identifies maxims of reflection, and principles of revision. This distinction is not relevant to my purposes here.

theoretical and practical reasoning that is best captured in terms of first-person vs. third person authority (Bagnoli 2013). This is because the constructivist view is concerned with granting authority to our claims, rather than with the objects or final ends of reasoning.

I favor a general account of reasoning, which is held together by principles. The appeal to principles is meant to guarantee not only the coherent structure of our thought, but also a coherent account of deliberation over time and of our personal relations.[54] The rational principles that guide our thought and action are principles that vulnerable, inter-dependent agents, placed in a similar perspective, endowed with similar emotional capacities, could share. The constraint that principles put on our practices makes a constitutive reference to others, not as representative of an idealized community, but as our co-legislators. Such principles are universal so as to ensure that we authoritatively respond to the demands of justification and requests of intelligibility pressed by others. Thus, ultimately, our revisions and changes are required to meet public standards of justification.[55] Both the need of objectivity and the means to achieve it are grounded on communal practices of validation and mutual recognition.

---

[54] Harman holds that the coherence principle is defeasible, Harman 1986, p. 11. On the contrary, I take it to be constitutive of reasoning. The difference is likely to be sharper than it appears because we also disagree about the status of the principle. On the Kantian view, coherence is not solely a logical principle, but also a principle of deliberation and reciprocity.

[55] My account shares the basic features of a Kantian account of public reason, see Bagnoli 2013. But there are different notions of public validation and justification. For instance, Baier claims that membership in a community is conventional and insists on the practice of "second-thinking". According to Baier, "the liberty to change one's mind importantly depends on the validation and recognition of others", Baier 1985, p. 66. I think Baier's notion of public validation is too weak to grant an interesting view of objectivity.

Bibliographical references

Bagnoli, C. (2007). "Phenomenology of the aftermath. Ethical theory and the intelligibility of moral experience". In S. Tenenbaum (Ed.), *New trends in moral psychology* (pp. 183–221). Amsterdam: Rodopi.

Bagnoli,C. (2011a). "Constructivism in meta-ethics". In E.N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (pp. 1–51).

Bagnoli,C. (2011b). "Moral perception and knowledge by principles. In J. Hernandez (Ed.), *New Intuitionism* (pp. 84–105). New York: Continuum.

Bagnoli, C. (2013). Constructivismabout practical knowledge. In C. Bagnoli (Ed.), Constructivism in ethics (pp. 153–182). Cambridge: Cambridge University Press.

Broome, J. (2004). *Weighing Lives*. Oxford: Oxford University Press.

Dancy, J. (1993). *Moral Reasons*. Oxford: Basil Blackwell.

Dancy, Jonathan (2004). *Ethics Without Principles*. Oxford University Press.

Dancy, J. (2006). *Ethics Without Principles*. Oxford: Oxford University Press.

Darwall, S. (2013). "Morality and Principle". In D. Bakhurst, M. Olivia Little & B. Hooker (eds.), *Thinking About Reasons: Themes From the Philosophy of Jonathan Dancy*. Oxford: Oxford University Press, 168-92.

Engstrom, S. (2013). "Constructivism and practical knowledge". In Bagnoli (ed.) *Constructivism in Ethics,* Cambridge: Cambridge University Press.

Hare, R. M. (1981). *Moral Thinking*. Oxford: Clarendon Press.

Hart, H.L.A. (1948-49). "The Ascription of Responsibilities and Rights". *Proceedings of the Aristotelian Society*, (49): 171-94. Reprinted in *Logic and Language*, (ed. by A. G. H. Flew, Oxford, 1955), 145-166.

Hart, H.L.A. (1961). *The Concept of Law*. Oxford: Clarendon Press.

Helm, P. (1968). "Defeasibility and Open Texture", *Analysis*, (28) No. 5, 173-175.

Herman, B. (1993). *The Practice of Moral Judgment*, Harvard University Press.

Holton, Richard (2002). "Principles and particularisms". *Proceedings of the Aristotelian Society* 76 (1):191-209.

Kant, I. *Kants gesammelte Schriften*, Preussische Akademie der Wissenschaften, Berlin, 1907. (Abbreviated Ak).

Kirchin, S. (2010). "The Shapelessness Hypothesis", *Philosophers' Imprint* 10(4): 1-28.

Korsgaard, C.M. (1996). *The Sources of Normativity.* Cambridge University Press.

Korsgaard, C.M. (2003). Realism and Constructivism in Twentieth-Century Moral Philosophy. Journal of Philosophical Research 28 (Supplement):99-122.

Korsgaard, C.M. (2008). *The Constitution of Agency: Essays on Practical Reason and Moral Psychology.* Oxford University Press.

Lance, M., M. Little (2008). "From particularism to defeasibility in ethics", in V. Strahovnik, M. Potrc & M. Norris Lance eds. 2008, *Challenging Moral Particularism*, Routledge, London. 53-74.

Lebar, Mark (2013). "Constructivism and particularism". In Bagnoli (ed.) *Constructivism in Ethics,* Cambridge: Cambridge University Press.

Little, M.O. (2000), "Moral Generalities Revisited", in *Moral Particularism*, B. Hooker & M. O. Little (eds.) Oxford: Clarendon Press, 276-304.

McDowell, J. (1981). "Non-cognitivism and rule-following". In S. Holtzman & Christopher M. Leich (eds.), *Wittgenstein: To Follow A Rule* . London: Routledge 141-62.

McKeever, S. & M. Ridge (2005). "What Does Holism Have to Do With Moral Particularism?," *Ratio*: 93-103.

McKeever, S. & M. Ridge (2006). *Principled Ethics: Generalism as a Regulative Ideal*, Oxford: Oxford University Press.

Millgram, Elijah (2005). *Ethics Done Right: Practical Reasoning as a Foundation for Moral Theory.* Cambridge University Press.

Nozick, R. (1968) "Moral Complications and moral structures", *American Journal of Jurisprudence* 13 (1): 1-50.

Nozick, R. (1981), Philosophical Explanations, Cambridge Mass: Belnknap of Harvard University Press.

O'Neill, Onora (1975). Acting on Principle: An Essay on Kantian Ethics. Columbia University Press.

O'Neill, O. (1989). *Constructions of Reason: Exploration of Kant's Practical Philosophy.* Cambridge: Cambridge University Press.

O'Neill, O. (1996). *Toward Justice and Virtue. A constructive account of practical reasoning.* Cambridge: Cambridge University Press.

O'Neill, O. (2015), *Constructing Authorities: Reason, Politics and Interpretation in Kant's Philosophy*. Cambridge: Cambridge University Press.

Pollock, J. (1986). "A Theory of Moral Reasoning," *Ethics* 96 (3), 506-523

Pollock, J. (1987). "Defeasible Reasoning". *Cognitive Science* 11 (4): 481-518.

Rawls, J. (1955). "Two concepts of rules". *Philosophical Review*, 64 (1):3-32.

Rawls, J. (1980). "Kantian constructivism in moral theory." *Journal of Philosophy*, 77 (9): 515-572.

Redondo, M.C. (2005). "Legal reasons: between universalism and particularism". *Journal of Moral Philosophy*, 2 (1): 47-68.

Redondo, M.C. (2012). Reasons for action and defeasibility. In J. Ferrer Beltrán & G.B. Ratti (eds.), *The Logic of Legal Requirements: Essays on Defeasibility*. Oxford University Press

Ross, W.D. *The Right and the Good*, Oxford: Oxford University Press,

Scanlon, T.M. (1998). *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Schroeder, M. (2011). "Holism, Weight, and Undercutting". *Noûs* 45 (2): 328 – 344.

Väyrynen, P. (2006). "Moral Generalism: Enjoy in Moderation", *Ethics*, 2006, pp. 707–41.

Väyrynen, P. (2009). A Theory of Hedged Moral Principles, in: Shafer-Landau (ed.) *Oxford Studies in Metaethics*, 91-132.

Wiggins, D. (1993). "A Neglected Position?", in *Reality, Representation, and Projection*, ed. J. Haldane & C. Wright (Oxford: Oxford University Press), 329-36.

Williams, B. (1963). "Ethical Consistency". *Proceeding of the Aristotelian Society* (39): 103-124. Reprinted in *Problems of the Self*, Cambridge University Press Williams, (1973): 166-186.

Williams, B. (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.